SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: GREENE, JOHN M
FLEISCHMANN, ROBERT D

(ii) TITLE OF INVENTION: HUMAN TUMOR NECROSIS FACTOR RECEPTOR

(iii) NUMBER OF SEQUENCES: 10

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C.
(B) STREET: 1100 NEW YORK AVENUE, NW, SUITE 600
(C) CITY: WASHINGTON
(D) STATE: DC
(E) COUNTRY: US
(F) ZIP: 20005-3934

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: US 08/469,637
(B) FILING DATE: 06-JUN-1995
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: STEFFE, ERIC K
(B) REGISTRATION NUMBER: 36,688
(C) REFERENCE/DOCKET NUMBER: 1488.0710001

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (202) 371-2600
(B) TELEFAX: (202) 371 2540

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1527 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 46..1248

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 46..106

(ix) FEATURE:
(A) NAME/KEY: mat_peptide
(B) LOCATION: 109..1248

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
CGCCCAGCCG CCGCCTCCAA GCCCCTGAGG TTTCCGGGGA CCACA ATG AAC AAG          54
                                                  Met Asn Lys
                                                  -21 -20

TTG CTG TGC TGC GCG CTC GTG TTT CTG GAC ATC TCC ATT AAG TGG ACC       102
Leu Leu Cys Cys Ala Leu Val Phe Leu Asp Ile Ser Ile Lys Trp Thr
            -15                 -10                 -5

ACC CAG GAA ACG TTT CCT CCA AAG TAC CTT CAT TAT GAC GAA GAA ACC       150
Thr Gln Glu Thr Phe Pro Pro Lys Tyr Leu His Tyr Asp Glu Glu Thr
          1               5                  10

TCT CAT CAG CTG TTG TGT GAC AAA TGT CCT CCT GGT ACC TAC CTA AAA       198
Ser His Gln Leu Leu Cys Asp Lys Cys Pro Pro Gly Thr Tyr Leu Lys
 15                  20                  25                  30

CAA CAC TGT ACA GCA AAG TGG AAG ACC GTG TGC GCC CCT TGC CCT GAC       246
Gln His Cys Thr Ala Lys Trp Lys Thr Val Cys Ala Pro Cys Pro Asp
                35                  40                  45

CAC TAC TAC ACA GAC AGC TGG CAC ACC AGT GAC GAG TGT CTA TAC TGC       294
His Tyr Tyr Thr Asp Ser Trp His Thr Ser Asp Glu Cys Leu Tyr Cys
            50                  55                  60

AGC CCC GTG TGC AAG GAG CTG CAG TAC GTC AAG CAG GAG TGC AAT CGC       342
Ser Pro Val Cys Lys Glu Leu Gln Tyr Val Lys Gln Glu Cys Asn Arg
          65                  70                  75

ACC CAC AAC CGC GTG TGC GAA TGC AAG GAA GGG CGC TAC CTT GAG ATA       390
Thr His Asn Arg Val Cys Glu Cys Lys Glu Gly Arg Tyr Leu Glu Ile
     80                  85                  90

GAG TTC TGC TTG AAA CAT AGG AGC TGC CCT CCT GGA TTT GGA GTG GTG       438
Glu Phe Cys Leu Lys His Arg Ser Cys Pro Pro Gly Phe Gly Val Val
 95                  100                 105                 110

CAA GCT GGA ACC CCA GAG CGA AAT ACA GTT TGC AAA AGA TGT CCA GAT       486
Gln Ala Gly Thr Pro Glu Arg Asn Thr Val Cys Lys Arg Cys Pro Asp
                115                 120                 125

GGG TTC TTC TCA AAT GAG ACG TCA TCT AAA GCA CCC TGT AGA AAA CAC       534
Gly Phe Phe Ser Asn Glu Thr Ser Ser Lys Ala Pro Cys Arg Lys His
            130                 135                 140
```


B

```
ACA AAT TGC AGT GTC TTT GGT CTC CTG CTA ACT CAG AAA GGA AAT GCA       582
Thr Asn Cys Ser Val Phe Gly Leu Leu Leu Thr Gln Lys Gly Asn Ala
        145                 150                 155

ACA CAC GAC AAC ATA TGT TCC GGA AAC AGT GAA TCA ACT CAA AAA TGT       630
Thr His Asp Asn Ile Cys Ser Gly Asn Ser Glu Ser Thr Gln Lys Cys
    160                 165                 170

GGA ATA GAT GTT ACC CTG TGT GAG GAG GCA TTC TTC AGG TTT GCT GTT       678
Gly Ile Asp Val Thr Leu Cys Glu Glu Ala Phe Phe Arg Phe Ala Val
175                 180                 185                 190

CCT ACA AAG TTT ACG CCT AAC TGG CTT AGT GTC TTG GTA GAC AAT TTG       726
Pro Thr Lys Phe Thr Pro Asn Trp Leu Ser Val Leu Val Asp Asn Leu
                195                 200                 205

CCT GGC ACC AAA GTA AAC GCA GAG AGT GTA GAG AGG ATA AAA CGG CAA       774
Pro Gly Thr Lys Val Asn Ala Glu Ser Val Glu Arg Ile Lys Arg Gln
            210                 215                 220

CAC AGC TCA CAA GAA CAG ACT TTC CAG CTG CTG AAG TTA TGG AAA CAT       822
His Ser Ser Gln Glu Gln Thr Phe Gln Leu Leu Lys Leu Trp Lys His
        225                 230                 235

CAA AAC AAA GAC CAA GAT ATA GTC AAG AAG ATC ATC CAA GAT ATT GAC       870
Gln Asn Lys Asp Gln Asp Ile Val Lys Lys Ile Ile Gln Asp Ile Asp
    240                 245                 250

CTC TGT GAA AAC AGC GTG CAG CGG CAC ATT GGA CAT GCT AAC CTC ACC       918
Leu Cys Glu Asn Ser Val Gln Arg His Ile Gly His Ala Asn Leu Thr
255                 260                 265                 270

TTC GAG CAG CTT CGT AGC TTG ATG GAA AGC TTA CCG GGA AAG AAA GTG       966
Phe Glu Gln Leu Arg Ser Leu Met Glu Ser Leu Pro Gly Lys Lys Val
                275                 280                 285

GGA GCA GAA GAC ATT GAA AAA ACA ATA AAG GCA TGC AAA CCC AGT GAC      1014
Gly Ala Glu Asp Ile Glu Lys Thr Ile Lys Ala Cys Lys Pro Ser Asp
            290                 295                 300

CAG ATC CTG AAG CTG CTC AGT TTG TGG CGA ATA AAA AAT GGC GAC CAA      1062
Gln Ile Leu Lys Leu Leu Ser Leu Trp Arg Ile Lys Asn Gly Asp Gln
        305                 310                 315

GAC ACC TTG AAG GGC CTA ATG CAC GCA CTA AAG CAC TCA AAG ACG TAC      1110
Asp Thr Leu Lys Gly Leu Met His Ala Leu Lys His Ser Lys Thr Tyr
    320                 325                 330

CAC TTT CCC AAA ACT GTC ACT CAG AGT CTA AAG AAG ACC ATC AGG TTC      1158
His Phe Pro Lys Thr Val Thr Gln Ser Leu Lys Lys Thr Ile Arg Phe
335                 340                 345                 350

CTT CAC AGC TTC ACA ATG TAC AAA TTG TAT CAG AAG TTA TTT TTA GAA      1206
Leu His Ser Phe Thr Met Tyr Lys Leu Tyr Gln Lys Leu Phe Leu Glu
                355                 360                 365
```


B

```
ATG ATA GGT AAC CAG GTC CAA TCA GTA AAA ATA AGC TGC TTA                1248
Met Ile Gly Asn Gln Val Gln Ser Val Lys Ile Ser Cys Leu
            370                 375                 380

TAACTGGAAA TGGCCATTGA GCTGTTTCCT CACAATTGGC GAGATCCCAT GGATGAGTAA     1308

ACTGTTTCTC AGGCACTTGA GGCTTTCAGT GATATCTTTC TCATTACCAG TGACTAATTT     1368

TGCCACAGGG TACTAAAAGA AACTATGATG TGGAGAAAGG ACTAACATCT CCTCCAATAA     1428

ACCCCAAATG GTTAATCCAA CTGTCAGATC TGGATCGTTA TCTACTGACT ATATTTTCCC     1488

TTATTACTGC TTGCAGTAAT TCAACTGGAA AAAAAAAAA                            1527
```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 401 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
Met Asn Lys Leu Leu Cys Cys Ala Leu Val Phe Leu Asp Ile Ser Ile
-21 -20                 -15                 -10

Lys Trp Thr Thr Gln Glu Thr Phe Pro Pro Lys Tyr Leu His Tyr Asp
 -5                   1                   5                  10

Glu Glu Thr Ser His Gln Leu Leu Cys Asp Lys Cys Pro Pro Gly Thr
             15                  20                  25

Tyr Leu Lys Gln His Cys Thr Ala Lys Trp Lys Thr Val Cys Ala Pro
         30                  35                  40

Cys Pro Asp His Tyr Tyr Thr Asp Ser Trp His Thr Ser Asp Glu Cys
     45                  50                  55

Leu Tyr Cys Ser Pro Val Cys Lys Glu Leu Gln Tyr Val Lys Gln Glu
 60                  65                  70                  75

Cys Asn Arg Thr His Asn Arg Val Cys Glu Cys Lys Glu Gly Arg Tyr
                 80                  85                  90

Leu Glu Ile Glu Phe Cys Leu Lys His Arg Ser Cys Pro Pro Gly Phe
             95                  100                 105

Gly Val Val Gln Ala Gly Thr Pro Glu Arg Asn Thr Val Cys Lys Arg
        110                 115                 120

Cys Pro Asp Gly Phe Phe Ser Asn Glu Thr Ser Ser Lys Ala Pro Cys
    125                 130                 135
```

```
Arg Lys His Thr Asn Cys Ser Val Phe Gly Leu Leu Leu Thr Gln Lys
140                 145                 150                 155

Gly Asn Ala Thr His Asp Asn Ile Cys Ser Gly Asn Ser Glu Ser Thr
                160                 165                 170

Gln Lys Cys Gly Ile Asp Val Thr Leu Cys Glu Glu Ala Phe Phe Arg
            175                 180                 185

Phe Ala Val Pro Thr Lys Phe Thr Pro Asn Trp Leu Ser Val Leu Val
        190                 195                 200

Asp Asn Leu Pro Gly Thr Lys Val Asn Ala Glu Ser Val Glu Arg Ile
    205                 210                 215

Lys Arg Gln His Ser Ser Gln Glu Gln Thr Phe Gln Leu Leu Lys Leu
220                 225                 230                 235

Trp Lys His Gln Asn Lys Asp Gln Asp Ile Val Lys Lys Ile Ile Gln
                240                 245                 250

Asp Ile Asp Leu Cys Glu Asn Ser Val Gln Arg His Ile Gly His Ala
            255                 260                 265

Asn Leu Thr Phe Glu Gln Leu Arg Ser Leu Met Glu Ser Leu Pro Gly
        270                 275                 280

Lys Lys Val Gly Ala Glu Asp Ile Glu Lys Thr Ile Lys Ala Cys Lys
    285                 290                 295

Pro Ser Asp Gln Ile Leu Lys Leu Leu Ser Leu Trp Arg Ile Lys Asn
300                 305                 310                 315

Gly Asp Gln Asp Thr Leu Lys Gly Leu Met His Ala Leu Lys His Ser
                320                 325                 330

Lys Thr Tyr His Phe Pro Lys Thr Val Thr Gln Ser Leu Lys Lys Thr
            335                 340                 345

Ile Arg Phe Leu His Ser Phe Thr Met Tyr Lys Leu Tyr Gln Lys Leu
        350                 355                 360

Phe Leu Glu Met Ile Gly Asn Gln Val Gln Ser Val Lys Ile Ser Cys
    365                 370                 375

Leu
380
```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 461 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

B

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```
Met Ala Pro Val Ala Val Trp Ala Ala Leu Ala Val Gly Leu Glu Leu
1               5                  10                  15

Trp Ala Ala Ala His Ala Leu Pro Ala Gln Val Ala Phe Thr Pro Tyr
            20                  25                  30

Ala Pro Glu Pro Gly Ser Thr Cys Arg Leu Arg Glu Tyr Tyr Asp Gln
        35                  40                  45

Thr Ala Gln Met Cys Cys Ser Lys Cys Ser Pro Gly Gln His Ala Lys
    50                  55                  60

Val Phe Cys Thr Lys Thr Ser Asp Thr Val Cys Asp Ser Cys Glu Asp
65                  70                  75                  80

Ser Thr Tyr Thr Gln Leu Trp Asn Trp Val Pro Glu Cys Leu Ser Cys
                85                  90                  95

Gly Ser Arg Cys Ser Ser Asp Gln Val Glu Thr Gln Ala Cys Thr Arg
            100                 105                 110

Glu Gln Asn Arg Ile Cys Thr Cys Arg Pro Gly Trp Tyr Cys Ala Leu
        115                 120                 125

Ser Lys Gln Glu Gly Cys Arg Leu Cys Ala Pro Leu Arg Lys Cys Arg
    130                 135                 140

Pro Gly Phe Gly Val Ala Arg Pro Gly Thr Glu Thr Ser Asp Val Val
145                 150                 155                 160

Cys Lys Pro Cys Ala Pro Gly Thr Phe Ser Asn Thr Thr Ser Ser Thr
                165                 170                 175

Asp Ile Cys Arg Pro His Gln Ile Cys Asn Val Val Ala Ile Pro Gly
            180                 185                 190

Asn Ala Ser Met Asp Ala Val Cys Thr Ser Thr Ser Pro Thr Arg Ser
        195                 200                 205

Met Ala Pro Gly Ala Val His Leu Pro Gln Pro Val Ser Thr Arg Ser
    210                 215                 220

Gln His Thr Gln Pro Thr Pro Glu Pro Ser Thr Ala Pro Ser Thr Ser
225                 230                 235                 240

Phe Leu Leu Pro Met Gly Pro Ser Pro Pro Ala Glu Gly Ser Thr Gly
                245                 250                 255
```


B

```
Asp Phe Ala Leu Pro Val Gly Leu Ile Val Gly Val Thr Ala Leu Gly
            260                 265                 270

Leu Leu Ile Ile Gly Val Val Asn Cys Val Ile Met Thr Gln Val Lys
        275                 280                 285

Lys Lys Pro Leu Cys Leu Gln Arg Glu Ala Lys Val Pro His Leu Pro
    290                 295                 300

Ala Asp Lys Ala Arg Gly Thr Gln Gly Pro Glu Gln Gln His Leu Leu
305                 310                 315                 320

Ile Thr Ala Pro Ser Ser Ser Ser Ser Ser Leu Glu Ser Ser Ala Ser
                325                 330                 335

Ala Leu Asp Arg Arg Ala Pro Thr Arg Asn Gln Pro Gln Ala Pro Gly
            340                 345                 350

Val Glu Ala Ser Gly Ala Gly Glu Ala Arg Ala Ser Thr Gly Ser Ser
        355                 360                 365

Asp Ser Ser Pro Gly Gly His Gly Thr Gln Val Asn Val Thr Cys Ile
    370                 375                 380

Val Asn Val Cys Ser Ser Ser Asp His Ser Ser Gln Cys Ser Ser Gln
385                 390                 395                 400

Ala Ser Ser Thr Met Gly Asp Thr Asp Ser Ser Pro Ser Glu Ser Pro
                405                 410                 415

Lys Asp Glu Gln Val Pro Phe Ser Lys Glu Glu Cys Ala Phe Arg Ser
            420                 425                 430

Gln Leu Glu Thr Pro Glu Thr Leu Leu Gly Ser Thr Glu Glu Lys Pro
        435                 440                 445

Leu Pro Leu Gly Val Pro Asp Ala Gly Met Lys Pro Ser
    450                 455                 460
```

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

GCCAGAGGAT CCGAAACGTT TCCTCCAAAG TAC 33

B

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

CGGCTTCTAG AATTACCTAT CATTTCTAAA AAT 33

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 31 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

GCGCGGATCC ATGAACAAGT TGCTGTGCTG C 31

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 34 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GCGCTCTAGA TTACCTATCA TTTCTAAAAA TAAC 34

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 31 base pairs

B

(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

GCGCGGTACC TCAGTGGTTT GGGCTCCTCC C 31

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 39 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

GCCAGAGGAT CCGCCACCAT GAACAAGTTG CTGTGCTGC 39

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 60 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

CGGCTTCTAG AATCAAGCGT AGTCTGGGAC GTCGTATGGG TACCTATCAT TTCTAAAAAT 60

```
      10                  30                  50
CGCCCAGCCGCCGCCTCCAAGCCCCTGAGGTTTCCGGGGACCACAATGAACAAGTTGCTG
                                                M  N  K  L  L
       70                  90                 110
TGCTGCGCGCTCGTGTTTCTGGACATCTCCATTAAGTGGACCACCCAGGAAACGTTTCCT
C  C  A  L  V  F  L  D  I  S  I  K  W  T  T  Q  E  T  F  P
      130                 150                 170
CCAAAGTACCTTCATTATGACGAAGAAACCTCTCATCAGCTGTTGTGTGACAAATGTCCT
P  K  Y  L  H  Y  D  E  E  T  S  H  Q  L  L  C  D  K  C  P
      190                 210                 230
CCTGGTACCTACCTAAAACAACACTGTACAGCAAAGTGGAAGACCGTGTGCGCCCCTTGC
P  G  T  Y  L  K  Q  H  C  T  A  K  W  K  T  V  C  A  P  C
      250                 270                 290
CCTGACCACTACTACACAGACAGCTGGCACACCAGTGACGAGTGTCTATACTGCAGCCCC
P  D  H  Y  Y  T  D  S  W  H  T  S  D  E  C  L  Y  C  S  P
      310                 330                 350
GTGTGCAAGGAGCTGCAGTACGTCAAGCAGGAGTGCAATCGCACCCACAACCGCGTGTGC
V  C  K  E  L  Q  Y  V  K  Q  E  C  N  R  T  H  N  R  V  C
      370                 390                 410
GAATGCAAGGAAGGGCGCTACCTTGAGATAGAGTTCTGCTTGAAACATAGGAGCTGCCCT
E  C  K  E  G  R  Y  L  E  I  E  F  C  L  K  H  R  S  C  P
      430                 450                 470
CCTGGATTTGGAGTGGTGCAAGCTGGAACCCCAGAGCGAAATACAGTTTGCAAAAGATGT
P  G  F  G  V  V  Q  A  G  T  P  E  R  N  T  V  C  K  R  C
      490                 510                 530
CCAGATGGGTTCTTCTCAAATGAGACGTCATCTAAAGCACCCTGTAGAAAACACACAAAT
P  D  G  F  F  S  N  E  T  S  S  K  A  P  C  R  K  H  T  N
      550                 570                 590
TGCAGTGTCTTTGGTCTCCTGCTAACTCAGAAAGGAAATGCAACACACGACAACATATGT
C  S  V  F  G  L  L  L  T  Q  K  G  N  A  T  H  D  N  I  C
      610                 630                 650
TCCGGAAACAGTGAATCAACTCAAAAATGTGGAATAGATGTTACCCTGTGTGAGGAGGCA
S  G  N  S  E  S  T  Q  K  C  G  I  D  V  T  L  C  E  E  A
      670                 690                 710
TTCTTCAGGTTTGCTGTTCCTACAAAGTTTACGCCTAACTGGCTTAGTGTCTTGGTAGAC
F  F  R  F  A  V  P  T  K  F  T  P  N  W  L  S  V  L  V  D
      730                 750                 770
AATTTGCCTGGCACCAAAGTAAACGCAGAGAGTGTAGAGAGGATAAAACGGCAACACAGC
N  L  P  G  T  K  V  N  A  E  S  V  E  R  I  K  R  Q  H  S
      790                 810                 830
TCACAAGAACAGACTTTCCAGCTGCTGAAGTTATGGAAACATCAAAACAAAGACCAAGAT
S  Q  E  Q  T  F  Q  L  L  K  L  W  K  H  Q  N  K  D  Q  D
      850                 870                 890
ATAGTCAAGAAGATCATCCAAGATATTGACCTCTGTGAAAACAGCGTGCAGCGGCACATT
I  V  K  K  I  I  Q  D  I  D  L  C  E  N  S  V  Q  R  H  I
      910                 930                 950
GGACATGCTAACCTCACCTTCGAGCAGCTTCGTAGCTTGATGGAAAGCTTACCGGGAAAG
G  H  A  N  L  T  F  E  Q  L  R  S  L  M  E  S  L  P  G  K
      970                 990                1010
AAAGTGGGAGCAGAAGACATTGAAAAAACAATAAAGGCATGCAAACCCAGTGACCAGATC
K  V  G  A  E  D  I  E  K  T  I  K  A  C  K  P  S  D  Q  I
     1030                1050                1070
CTGAAGCTGCTCAGTTTGTGGCGAATAAAAAATGGCGACCAAGACACCTTGAAGGGCCTA
L  K  L  L  S  L  W  R  I  K  N  G  D  Q  D  T  L  K  G  L
     1090                1110                1130
ATGCACGCACTAAAGCACTCAAAGACGTACCACTTTCCCAAAACTGTCACTCAGAGTCTA
```

FIGURE 1(A)

```
M  H  A  L  K  H  S  K  T  Y  H  F  P  K  T  V  T  Q  S  L
      1150                1170                1190
AAGAAGACCATCAGGTTCCTTCACAGCTTCACAATGTACAAATTGTATCAGAAGTTATTT
K  K  T  I  R  F  L  H  S  F  T  M  Y  K  L  Y  Q  K  L  F
      1210                1230                1250
TTAGAAATGATAGGTAACCAGGTCCAATCAGTAAAAATAAGCTGCTTATAACTGGAAATG
L  E  M  I  G  N  Q  V  Q  S  V  K  I  S  C  L  *
      1270                1290                1310
GCCATTGAGCTGTTTCCTCACAATTGGCGAGATCCCATGGATGAGTAAACTGTTTCTCAG
      1330                1350                1370
GCACTTGAGGCTTTCAGTGATATCTTTCTCATTACCAGTGACTAATTTTGCCACAGGGTA
      1390                1410                1430
CTAAAAGAAACTATGATGTGGAGAAAGGACTAACATCTCCTCCAATAAACCCCAAATGGT
      1450                1470                1490
TAATCCAACTGTCAGATCTGGATCGTTATCTACTGACTATATTTTCCCTTATTACTGCTT
      1510
GCAGTAATTCAACTGGAAAAAAAAAAA
```

FIGURE 1(B)

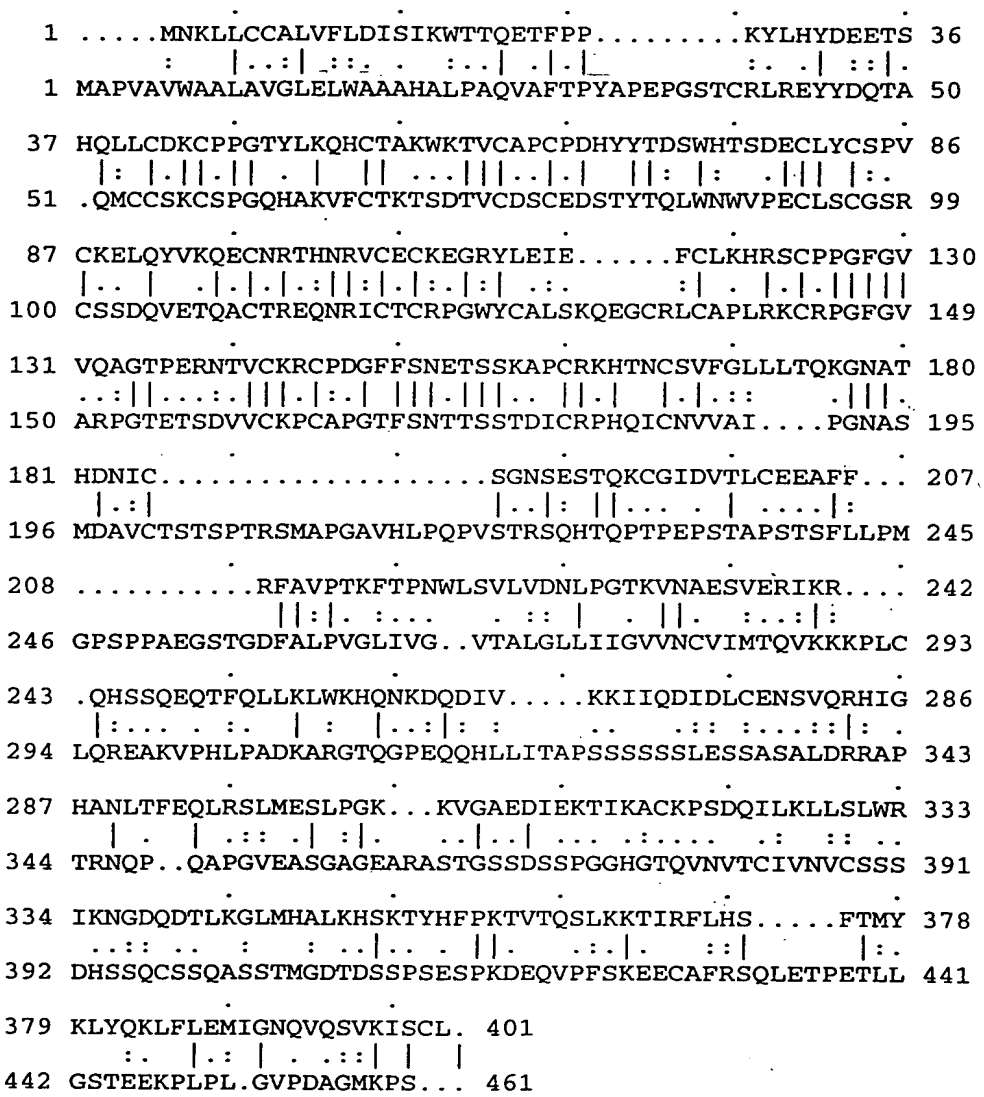
1 .....MNKLLCCALVFLDISIKWTTQETFPP.....KYLHYDEETS 36
1 MAPVAVWAALAVGLELWAAAHALPAQVAFTPYAPEPGSTCRLREYYDQTA 50
37 HQLLCDKCPPGTYLKQHCTAKWKTVCAPCPDHYYTDSWHTSDECLYCSPV 86
51 .QMCCSKCSPGQHAKVFCTKTSDTVCDSCEDSTYTQLWNWVPECLSCGSR 99
87 CKELQYVKQECNRTHNRVCECKEGRYLEIE.....FCLKHRSCPPGFGV 130
100 CSSDQVETQACTREQNRICTCRPGWYCALSKQEGCRLCAPLRKCRPGFGV 149
131 VQAGTPERNTVCKRCPDGFFSNETSSKAPCRKHTNCSVFGLLLTQKGNAT 180
150 ARPGTETSDVVCKPCAPGTFSNTTSSTDICRPHQICNVVAI....PGNAS 195
181 HDNIC.....SGNSESTQKCGIDVTLCEEAFF... 207
196 MDAVCTSTSPTRSMAPGAVHLPQPVSTRSQHTQPTPEPSTAPSTSFLLPM 245
208 .....RFAVPTKFTPNWLSVLVDNLPGTKVNAESVERIKR.... 242
246 GPSPPAEGSTGDFALPVGLIVG..VTALGLLIIGVVNCVIMTQVKKKPLC 293
243 .QHSSQEQTFQLLKLWKHQNKDQDIV.....KKIIQDIDLCENSVQRHIG 286
294 LQREAKVPHLPADKARGTQGPEQQHLLITAPSSSSSSLESSASALDRRAP 343
287 HANLTFEQLRSLMESLPGK...KVGAEDIEKTIKACKPSDQILKLLSLWR 333
344 TRNQP..QAPGVEASGAGEARASTGSSDSSPGGHGTQVNVTCIVNVCSSS 391
334 IKNGDQDTLKGLMHALKHSKTYHFPKTVTQSLKKTIRFLHS.....FTMY 378
392 DHSSQCSSQASSTMGDTDSSPSESPKDEQVPFSKEECAFRSQLETPETLL 441
379 KLYQKLFLEMIGNQVQSVKISCL. 401
442 GSTEEKPLPL.GVPDAGMKPS... 461

FIGURE 2